The Hereditary Abilities Study: Hereditary Components in a Psychological Test Battery

STEVEN G. VANDENBERG[1]

Institute of Human Biology, University of Michigan, Ann Arbor

FROM 1952 UNTIL 1957 an interdisciplinary study of twins was conducted at the University of Michigan in the Institute of Human Biology under the direction of Lee R. Dice. The general design of the study, the selection of twins, and the diagnosis of zygosity are described elsewhere (Sutton, Vandenberg, and Clark, 1962). This paper will present selected results obtained on the battery of psychological tests.

A total of eighty-two pairs of like-sexed twins were examined. Of these, forty-five pairs were judged to be identical and thirty-seven pairs fraternal twins on the basis of bloodtyping. The twins were of high school age and came from schools in Ann Arbor, Ypsilanti, Dearborn, and Detroit. Because other reports will be published no detailed information about each psychological test will be given here, other than a brief statement concerning the nature of the trait it attempts to measure. The author holds no brief for the validity of these statements, which have been taken from descriptions by the various test authors. For the information of investigators desirous of repeating or expanding some of these findings, the time required for instructions and the actual taking of each test are indicated in the statistical tables.

METHOD

The importance of the hereditary component in these measurements was studied by an analysis of variance technique. The within-pair variance of the fraternal or dizygous (DZ) twins was divided by the within-pair variance of the identical or monozygous (MZ) twins. The statistical significance of this ratio may be evaluated by Fisher's F test.

To allow comparisons with previous studies, Holzinger's h^2 measure of heritability was also calculated, in spite of the frequent misinterpretations of this index. Clark (1956) has given the correct formula for this calculation which does not assume equal variance for the identical and fraternal twins.

COGNITIVE ABILITIES

The establishment of the Intelligence Quotient as an index of the general level of mental ability was a great scientific achievement with far-reaching

---

Received June 22, 1961.

Supported in part by MacGregor Fund of Detroit and by grants M-1045 and RG-5527 of the National Institute of Mental Health.

[1]Present address: Longitudinal Twin Study, Department of Pediatrics, University of Louisville, Louisville, Ky.

practical effects in education, psychiatry, law, etc., because it made it possible to distinguish lack of ability from lack of interest, mental illness, or restricted education. The amazing success of this concept was perhaps partly due to the fact that it was an overly simple one. Wide variations in mental abilities can, and do, exist within any single individual.

It is, for instance, rather common to find someone who is highly gifted verbally but who is poor in numerical ability or in the ability to understand mechanical principles and vice versa. One overall index number can hardly be expected to provide an accurate measure of all the different mental abilities. Guilford (1960) has estimated that there may be as many as one hundred and twenty separate abilities and that at present there are perhaps fifty-five known, admittedly of varying social importance. For this reason it is a gross oversimplification to speak of the heritability of "intelligence." The hereditary component will vary from test to test. Therefore, it becomes necessary to study the hereditary component of a number of separate cognitive abilities.

There is no agreement as yet between psychologists on the number or precise nature of these separate abilities. For a critical review of various theories see Loevinger (1951). French (1951) has prepared a list of the psychological tests which measure the better established cognitive factors. Guilford (1956) proposed an integral theory about the structure of intelligence in which the various cognitive factors are presented as fitting into the cells of a three-dimensional schema based on three principles of organization, namely (1) the nature of the mental operations to be performed (i.e., recognition, analogy reasoning, memory, etc.) (2) the content of the test (i.e., verbal, pictoral, etc.), and (3) the general nature of the answer or "product" which is desired (i.e., relations, transformations, classes, etc.)

No systematic attempt was made to cover all types of cognitive abilities, but a wide variety of tests was used, consisting of both individually administered verbal or performance tests and group-administered paper-and-pencil tests, in an attempt to "sample" various types of tests said to be practically uncorrelated with each other.

The Hereditary Abilities Study was planned in 1952. At that time it was thought best to make the Chicago Primary Mental Abilities Tests of L. L. Thurstone (1938, 1941-a and b) the core of the battery, and to add other tests that claimed to be measuring abilities which seemed to be rather different from, and uncorrelated with, those measured by this test battery.

In some paper-pencil mental tests the subject is required to select the proper answer from several alternatives. Because in such tests he may obtain some correct answers by chance, a correction for guessing is often used to obtain the achievement score. The correction usually consists of subtracting from the total number of right answers the number of wrong answers divided by the number of alternatives provided. Whenever such a correction has been applied, the column labeled "scoring formula" in tables 1 to 3 shows the achievement score used in the calculations.

Table 1 presents the heritability indices and F ratios for the seventeen subtests of the Primary Mental Abilities Test battery. Three of the F ratios are significant beyond the .01 level of probability and another eight are significant be-

TABLE 1. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS FOR THE SEVENTEEN PARTS OF THE THURSTONE PRIMARY MENTAL ABILITIES (PMA) TESTS OF N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

Name of test	Scoring formula	h^2	F	p	N_{dz}	N_{mz}	Time required
Number							
1. Addition	R-W	58*	2.40	.01	37	45	9
2. Multiplication	R-W	43	1.75	.05	37	45	8
3. Three higher	R-W	51	2.05	.05	37	45	11
Verbal							
4. Sentences	R	45	1.80	.05	37	45	8
5. Vocabulary	R	00	.99		37	45	7
6. Completion	R	67	3.01	.01	37	45	9
Spatial							
7. Flags	R-W	34	1.80	.05	36	45	15
8. Figures	R-W	27	1.79	.05	36	45	11
9. Cards	R-W	52	1.77	.05	36	45	11
Word fluency							
10. First Letters	R	47	1.90	.05	35	44	8
11. Four letter words	R	37	1.59		35	44	7
12. Suffixes	R	57	2.31	.01	35	44	7
Reasoning							
13. Letter series	R	07	1.08		37	45	12
14. Letter grouping	R	44	1.78	.05	37	45	11
15. Pedigrees	R	22	1.28		37	45	11
Memory							
16. First names	R	48	1.57		34	45	10
17. Word-number	R	60	1.05		34	45	10

* Decimal points are omitted in all h^2 results.

Total scores for six primary abilities in PMA battery:

Name of test	h^2	F	p	N_{dz}	N_{mz}	Time required
Number	61	2.58	.01	37	45	28
Verbal	62	2.65	.01	37	45	24
Spatial	59	1.77	.05	46	45	37
Word fluency	61	2.57	.01	35	44	22
Reasoning	28	1.40		37	45	34
Memory	20	1.26		34	45	20

yond the .05 level. When scores on the subtests are added to give measures of the six primary abilities, we find that three F ratios are beyond the .01 level and one additional F ratio is beyond the .05 level of significance. The F ratios for the tests measuring reasoning and memory are not statistically significant. It is perhaps relevant here that many studies have shown that both reasoning and memory are phenomena which are too complex to be properly considered single cognitive factors. It is possible that the performance on the reasoning and memory tests used in this study depended at times on a different combination of factors for one twin partner than for the other. Under such conditions no significant heritability estimate would be obtained even though heritability estimates of the separate factors would perhaps be significant, if measures of the separate factors could be obtained. The type of material and the method of presentation have also been found to make important differences in the results of such tests. Further studies of heritability should take such findings into account.

Various other cognitive tests were added to the battery, largely on the basis of their frequent use here or in Great Britain in clinic, school, or hospital, or because they seemed to have some unique feature that might make interesting their inclusion in this type of study. A few of these tests are more commonly called achievement tests, since they are concerned with the innate potential upon which education has worked and which together with differences in environment and education produces the individual differences we observe.

The additional cognitive tests selected were these:

18. The vocabulary test of the Wechsler Intelligence Scale for Children (1949).
19. An abbreviated form of the Full Range Picture Vocabulary Test constructed by R. B. Ammons, (1949-a and b; 1950-a, and c).
20. The Gray Oral Reading Paragraphs (1915).
21. The Spelling Test, Form D, a part of the Stanford Achievement Test, constructed by Kelley, Ruch and Terman (1950). The even-numbered items from the complete list of one hundred words were used.
22. The Spelling Test of the Differential Aptitude Test constructed by Bennett, Seashore and Wesman (1947), a true-false test.
23. The Space Relations Test of the Differential Aptitude Test.
24. The Object Aperture Test, an experimental test of DuBois and Gleser (1948), another space visualization test.
25. The Survey of Mechanical Insight Test, constructed by Miller (1945).
26. The Progressive Matrices Test by J. C. Raven (1938 and 1940), said to be a measure of a general intellectual factor "g" rather than of any specialized intellectual ability.
27. Arthur Stencils, Form 1, by Grace Arthur (1944).
28. Block Designs Test, originally developed by Kohs (1923). We used the adaptation which forms one of the eleven subtests of the Bellevue Intelligence Scale by Wechsler (1944).
29. The Alexander Passalong Test forms one of the subtests developed by Alexander (1935) for his battery of tests of "concrete" intelligence, which he believed to be related to mechanical aptitude.
30. Cube Construction Test. This is one of the subtests of the Cornell-Coxe Point Scale of Intelligence (1934). Two scores were calculated for this test. In the first one, speed alone was counted, provided the solutions were correct.
31. In the second score partial credit was given for solutions which were not completely correct so that the total score was a combination of accuracy and speed.

32, 33. Parts 1 and 2 of the Heim AH4 Test. This test was constructed by Alice Heim (1947, 1955) as a measure of "general" intelligence. It is also called "omnibus" test, i.e., it consists of several kinds of items which are interspersed so that the score depends on an unknown mixture of the performances on these different types of items.

34. The Memory for Faces. This test forms part of the George Washington Test of Social Intelligence developed by F. A. Moss, Thelma Hunt and K. T. Omwake (1930).
35. The Memory for Digits Subtest of the Wechsler Intelligence Scale for Children (1949).

The standard instructions furnished with these tests were used on all occasions. Table 2 shows the heritability indices, F ratios and significance levels of the additional cognitive tests.

For three of these eighteen additional cognitive tests the F ratios exceed the .01 level of significance, and for another four the F ratios exceed the .05 level of significance. There is evidence that the method of examination influences heritability estimates in the fact that the Differential Aptitudes Test (DAT) of

TABLE 2. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS FOR VARIOUS COGNITIVE AND ACHIEVEMENT TESTS OF N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

	Name of test	Scoring formula	h^2	F	p	N dz	N mz	Time required
18.	WISC Vocabulary	As usual	54	2.19	.05	31	42	10
19.	Ammons Picture Vocabulary (abbrev.)	R	13	1.42		35	45	10
20.	Gray Oral Reading Paragraphs	Speed and accuracy	49	1.97	.05	36	45	10
21.	Stanford Spelling Achievement Test	R	60	2.47	.01	33	43	20
22.	DAT Spelling Test	R	40	1.67		34	44	5
23.	DAT Space Relations Test	R	02	1.02		34	44	12
24.	DuBois Object Aperture, Form A	R	17	1.21		34	45	10
25.	Miller Mechanical Insight	R	31	1.44		34	45	20
26.	Raven Progressive Matrices	R	44	1.80	.05	33	43	20
27.	Arthur Stencils, Form 1	Speed and accuracy	25	1.34		34	42	20
28.	Kohs' Blocks	Speed and accuracy	60	2.53	.01	34	45	10
29.	Alexander Passalong Test	Speed and accuracy	13	1.16		32	38	20
30.	Coxe Cube Construction	Speed	33	1.50		35	41	10
31.	Coxe Cube Construction	Speed and accuracy	21	1.27		35	41	10
32.	Heim, AH4, Part 1	R	55	2.20	.01	34	45	12
33.	Heim, AH4, Part 2	R	36	1.56		34	45	12
34.	Hunt Memory for Faces	R	00	.98		34	45	11
35.	WISC Memory for Digits	As usual	46	1.85	.05	31	42	10

spelling, a true-false type objective test, has an F ratio of only 1.67, which is not significant, while the Stanford Spelling Achievement Test, which is a dictation test in which the subject has to write the words himself, had an F ratio of 2.47, significant at the .01 level. Similarly, in the Vocabulary Test of the Wechsler Intelligence Scale for Children (WISC) which calls for definitions by the subject, the F ratio was a highly significant 2.19, while the Vocabulary Subtest of the Primary Mental Abilities (PMA) battery, which is a multiple-choice type objective test, had a nonsignificant F ratio of .99; and the abbreviated form of the Ammons Picture Vocabulary Test, which calls for identification of the one picture out of four which illustrates a given word, had an F ratio of 1.42, which is also nonsignificant.

PERCEPTUAL INTEGRATION

Among the tests believed to be measuring abilities considerably different from those measured by the PMA battery and the other cognitive tests were a number of experimental tests of perception, many of which were constructed in Thurstone's Psychometric Laboratory then located at the University of Chicago.[1] These tests are numbered 34 through 48 in this paper. In most of these tests the subject has to impose a perceptual reorganization on the material if he is to perceive "meaning" and perform well in them. Thurstone called the hypothetical factor common to tests requiring this type of ability the "closure" factor

[1]The Psychometric Laboratory moved to the University of North Carolina in 1952. Lyle V. Jones is its current director.

and attributed considerable importance to this factor in his thinking about mental abilities and personality traits.

These tests were chosen from among the many developed by Thurstone and his associates:

36. Concealed Figures. This test is similar to the Gottschalk Embedded Figures Test. In it the subject has to indicate whether a geometrical design is imbedded in each of four other designs.
37. Copying Designs. The subject has to connect the correct dots in a 4 x 8 matrix of dots.
38. Figure Classification. This test calls for the formation of a concept which will separate figures into two classes.
39. Hidden Four Letter Words. The subject has to locate and circle four letter words which occur in a page of continuous letters.

40, 41. Gestalt Completion A and B calls especially on the above named "closure" ability.

42. Identical Forms. This test is considered a rather pure measure of simple perceptual speed.
43. Identical Numbers. The subject has to circle the number 217 whenever it occurs in a page of three digit numbers. This is a measure of clerical speed and accuracy.
44. Incomplete Words. In this test the subject has to supply one or more missing letters to each word.
45. Reading Mirror Script.

46, 47. Mutilated Words A & B. The subject has to write down as many of the words as he is able to recognize. This test also calls for the above named "closure" aptitude.

Most of Thurstone's tests were used in his factor analysis of perceptual tests (1949), but a few are unpublished tests, developed since that time and kindly furnished by him for use in the Michigan Twin Study. The other perceptual tests used were these:

48. The McGill Closure Test by Mooney (1954), similar to the Gestalt Completion Tests of Thurstone.
49. Perception of Upright. A rod and frame apparatus similar to the one described by Witkin (1954) was used to measure the subject's ability to determine the true vertical in spite of a distorting field. The task consists of turning a knob to align a rod so that it is truely upright even though it is surrounded by a square frame which is tilted to one side or the other. Both rod and frame are painted with luminous paint and the subject performs this task in a room which is completely darkened. The score was the average absolute deviation from true vertical for sixteen trials with the frame at eight different degrees of tilt to the right or the left on each trial.
50. Size Constancy Experiment. The apparatus was similar to the one described by Thurstone (1949). The task consists of matching the size of a variable triangle with the size of a triangle of fixed size.
51. Critical Flicker Fusion. The maximum frequency at which a subject could still see that a light was flickering was determined with an electronic apparatus constructed by Dr. Lee R. Dice. The design of the apparatus controlled for the subjective brightness level of the flickering light and delivered a rectangularly shaped on and off pattern of stimulation. While the apparatus permitted varying the light-dark ratio, it was decided to use the standard 1 to 1 ratio throughout the experiments.

Table 3 presents the heritability indices and F ratios for these seventeen perceptual tests, of which fourteen were of paper-and-pencil variety. Three of the F ratios exceed the .01 level of significance and another five exceed the .05 level

TABLE 3. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS FOR SEVENTEEN PERCEPTUAL TESTS OF N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

Name of test	Scoring formula	h^2	F	p	N dz	N mz	Time required
36. Concealed Figures (Thurstone)	R-W	67	3.05	.01	34	43	14
37. Copying Designs (Thurstone)	R	33	1.48		34	44	7
38. Figure Classification (Thurstone)	R-W	35	1.58		34	43	16
39. Hidden 4 Letter Words (Thurstone)	R	52	2.10	.05	35	45	8
40. Gestalt Completion Form A, (Thurstone)	R	39	1.63		36	45	5
41. Gestalt Completion Form B, (Thurstone)	R	57	2.33	.01	34	45	5
42. Identical Forms (Thurstone)	R-W/4	38	1.60		34	45	7
43. Identical Numbers (Thurstone)	R	34	1.52		35	45	5
44. Incomplete Words (Thurstone)	R	45	1.81	.05	36	44	6
45. Reading Mirror Script (Thurstone)	R-W/3	46	1.86	.05	33	45	7
46. Mutilated Words, Form A (Thurstone)	R	27	1.37		36	44	6
47. Mutilated Words, Form B (Thurstone)	R	50	2.00	.05	34	45	6
48. McGill Closure Test (Mooney)	R	57	2.31	.01	37	45	6
49. Perception of Upright (Witkin)	Degrees of tilt	40	1.69		32	41	20
50. Size Constancy (Thurstone)	Error in cm	47	1.89	.05	33	38	55
51. Critical Flicker Frequency (Dice)	Mean value	44	1.77		25	29	20

of significance. That the heritability index is influenced by the kind of items which form the test as well as by the method of presentation is further demonstrated by the fact that the results were different for both tests for which two alternate forms were used. The mean number of items answered correctly for Forms A and B of the Gestalt Completion Test was 14.27 and 16.49; and for the Mutilated Words Test, Forms A and B, the means were 16.42 and 21.84. These tests were always administered in the same order: Form A first, followed several tests later by Form B. Other tests were interspersed to maintain the subject's interest by the variety of tasks. We cannot say whether these differences in the means are due to differences in difficulty level of the tests, or whether they are due to the fact that subjects had in a sense obtained practice while taking Form A, so that they might be expected to do better on Form B. In any event, the heritability was greater (that is, the hereditary control was more clearly established) for those alternate forms of the tests on which the subjects performed better on the average.

Some of the tests listed in table 3 as perceptual tests might perhaps be listed equally well in table 6 with the personality tests, since the way in which an individual interprets the sensory input received has been shown to be determined in part by his personality structure. This is the rationale behind the well known Rorschach Test, in which the subject is asked to interpret random ink blots.

MOTOR SKILLS

Tests of various motor abilities were included in the study for various reasons. First of all, there is the popular interest in the question "whether athletic ability is inherited." While the tests used are not measuring athletic ability in any particular sport, they may be related to it insofar as they measure some basic types of motor coordination required for many kinds of athletic ability. Another reason for inclusion of these tests lies in the fact that the majority of items in tests of "intelligence" for babies and infants are concerned with motor development. This is based on a recently questioned hypothesis that individual differences in motor skills may have something to do with the development of subsequent individual differences in intellectual skills. Finally, there is some evidence in studies by Eysenck (1952) and others that impairment of the speed and smoothness of coordination of motor performance may be related to personality traits such as neuroticism or, more generally, tenseness, when there is no direct organic basis for such impairment.

These tests of motor skills were selected:

52. Mirror Drawing: The apparatus sold by the Lafayette Instrument Company of Lafayette, Indiana was used. The subject has to trace a pattern while watching his work in a mirror. The pattern used was a six-pointed star formed by two parallel lines. A score based on the speed and accuracy of the performance was obtained for the right and also for the left hand.

53. Tweezer Dexterity. In this test designed by Johnson O'Connor (1928, 1938) the subject uses tweezers to place small brass pins into holes drilled into a board.

54. Santa Ana Dexterity Test. This test is described in the volume on the apparatus tests in the Army Air Force Psychology Program during World War II (Melton, 1947). The subject is asked to turn pegs in a board 180 degrees. There were three trials of 30 seconds each, and the right and left hand scores were the averages for the three trials.

55. Hand Steadiness. The apparatus sold by the Lafayette Company was used. The subject has to insert a stylus for a distance of 1 cm. into each of nine holes without touching the sides. The holes have decreasing diameters. A buzzer and dry cell battery were connected to the apparatus in such a way that a soft buzzing sound was made when the side of the hole was touched. Three trials were permitted for each hole, and 3, 2, or 1 points were earned, depending on whether correct insertion occured on the first, second or third trial. After failure on two consecutive holes the subject was stopped. The score was the total number of points for the nine holes.

56. Rotary Pursuit. The apparatus sold by the Lafayette Company was used. A small circular brass spot is located near the circumference of a record turntable. The subject's task is to try to keep a metal stylus on this moving brass spot. As long as he does a circuit is closed which activates a clock. The circuit is broken when the turntable motor stops, as it does for regular, built-in, rest periods between the one minute trials. A buzzer provided a faint noise during successful performance. Three trials for each hand were given in alternating order, so that "central" learning effects, i.e., which were not specific to each hand, would be distributed evenly.

57. Card Sorting. The rack sold by the Lafayette Company was used with Rook playing cards manufactured by the Parker Company. These cards come in decks without jack, queen, king or ace. Instead there are cards with one, eleven, twelve, or thirteen spots. The subject's task consists of sorting these cards, one by one, into pockets in the rack labeled 1 to 13 in the way postmen sort mail.

58. Continuous Mazes. The maze published in the laboratory exercises of MacKinnon and Henle (1948) was used. It consists of twenty-four little mazes, connected one after another, so that the performance on the total maze can be scored easily in terms of the number of smaller mazes completed. The subjects were given 3 minutes for this task. This time limit was chosen with the expectation, which proved correct, that nobody would obtain a perfect score.

59. Beam Balancing. In this test the subject was asked to walk on a 1 inch thick board raised 1 foot from the floor. The subjects were asked to walk in a heel-to toe-fashion. The board was 10 feet-long and marked in half-feet sections, so that it was easy to note at what point the subjects lost their balance.

60. Body Sway. The Ataxiameter constructed by Miles (1950) was used. This consists of a square frame suspended above the subject who wears a light head piece to which four threads are attached which record his movements in four directions. The score was the sum of the movements during a 60 second period. The subject was not asked to close his eyes, nor were any suggestions about swaying or falling given to the subject.

61. Average Reaction Time. The subject had to react to a light by pressing a telegraph key. Two series of twelve trials were run. In the first series the light was preceded by a buzzer at a fixed 2 second interval, while during the second series the warning buzzer preceded the light which formed the stimulus to which he had to react by a period which varied from 1 second to 5 seconds. The score for variable 61 was the mean reaction time for all twenty-four trials.

62. Difference between the reaction time on the series with regular and the series with the irregular warning period. Because it is well known that most subjects quickly learn to estimate the interval between warning signal and stimulus well enough to react almost instantaneously, the variable warning interval is a necessary condition for any longer series of measures of a true reaction time.

63. The variability around the mean of the reaction time for the trials with the fixed time between warning and stimulus.

64. The variability around the mean of the trials with a changing interval.

65. The ratio of these two variabilities. The rationale for this variable was the same as for variable 62. The results were again negative.

Table 4 presents the results of these motor skill tests. Of the fourteen tests reported here, six reached an F-value which exceeded the .01 level of significance, and two more were beyond the .05 level of significance. An additional piece of information which was obtained rather incidentally was that there were differences between the heritability of the same skill when performed with the right and with the left hand. Table 5 presents a comparison of the results obtained for the two hands. There is progression from fine hand movements towards gross hand-arm movements, but no clear indication emerges from the results to indicate that the degree of heritability is related to the type of movement required. Nor is any relation evident to the degree of familiarity with the type of activity required, although a minor hypothesis entertained during the selection was that unfamiliar tasks would show greater genetic differences, since the environmental factor would be minimized.

PERSONALITY TRAITS

Most of the personality tests selected for this study are subject to all the criticisms leveled at self-report questionnaires: they are subject to faking, they depend too much on the subject's limited understanding and knowledge of his own behavior, and they do perhaps not touch on underlying dynamics of his behavior. Limitations of time however required that the main effort be concen-

TABLE 4. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS FOR THE MOTOR SKILLS TESTS FOR N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

Name of test	Scoring formula	h^2	F	p	N dz	N mz	Time required
52. Mirror Drawing	right hand	70	3.38	.01	32	42	5-15
	left hand	24	1.31		35	41	5-15
53. Tweezer Dexterity	right hand	71	3.40	.01	32	43	5
	left hand	63	2.73	.01	32	43	5
54. Santa Ana Dexterity	right hand	58	2.41	.01	34	40	5
	left hand	05	1.05		34	40	5
55. Hand Steadiness	right hand	37	1.59		35	44	5
	left hand	17	1.21		35	44	5
56. Rotary Pursuit	right hand	52	2.08	.05	32	43	5
	left hand	32	1.46		32	43	5
57. Card Sorting	right hand	61	2.57	.01	34	41	10
	left hand	71	3.42	.01	34	41	10
58. Continuous Mazes	Speed and accuracy	20	1.24		34	45	4
59. Beam Balancing (Seashore)	shoes on, hands behind back	48	1.94	.05	32	41	10
60. Body Sway (Miles)		24	1.32		30	40	15
61. Mean Reaction Time	Regular interval	22	1.29		29	43	20
	irregular interval	—	.76		29	43	20
62. Difference between Regular and Irregular Intervals		—	.89		29	43	No extra time needed
63. Variability (SD_1)	regular interval	28	1.41		29	43	
64. Variability (SD_2)	irregular interval	04	1.05		29	43	
65. Ratio of SD_1 to SD_2		—	.97		29	43	

TABLE 5. COMPARISON OF F RATIOS FOR RIGHT AND LEFT HAND PERFORMANCE

Name of test	RH	p	LH	p
Mirror Drawing	3.38	.01	1.31	
Tweezer Dexterity	3.40	.01	2.73	.01
Santa Ana Dexterity	2.41	.01	1.05	
Hand Steadiness	1.59		1.21	
Rotary Pursuit	2.08		1.46	
Card Sorting (two-hand task)	2.57	.01	3.42	.01
Dart Throwing	.63		1.23	

trated on tests which could be group administered. Three of the five tests reported here were of the paper-and-pencil variety, but they furnished a total of twenty scores. The paper and pencil tests chosen were the Thurstone Temperament Schedule which provides seven scores:

66. Active
67. Vigorous
68. Impulsive
69. Dominant
70. Stable
71. Sociable
72. Reflective

and the Cattell (1950) Junior Personality Quiz which provides twelve scores:

73. Emotional sensitivity
74. Nervous tension
75. Will control
76. Neuroticism
77. Impatient dominance
78. Cyclothymia
79. Adventurous cyclothymia
80. Socialized morale
81. Independent dominance
82. Energetic conformity
83. Surgency
84. Intelligence

The scale for intelligence is rather short and did perhaps not provide enough of a challenge to our subjects which might explain the low heritability for these subjects.

The third paper-and-pencil test was:

85. The Maudsley Word-Connection Test constructed by Crown (1952) in an attempt to modify Jung's well known word association technique for use in group administration. For each of the words, the subject is given a choice of two associations, one a rather common one and one a less common one which the author had observed in his work with disturbed subjects. The score is the number of uncommon choices.

The last paper-and-pencil test was 87, a printed version of the Stroop Color Naming Test to be described here. The remaining two tests, which were "Performance" tests of personality, furnished seven scores.

86. The Stroop Color Naming Test consists of two parts. In the first part the subject is asked to name the colors of series of circular dots colored red, blue, yellow or green. This furnishes an average color-naming time for the dots. In the second part the subject is presented with the words RED, YELLOW, BLUE, and GREEN repeated in random order on an 8 1/2 by 11 sheet. These words are painted in the four colors, but the word RED may be painted blue, yellow or green. The subject's task is to name the color of the paint rather than to read the word. The inhibition of the familiar response based on reading the word delays appreciably the naming of the color of the words. This delay is measured by the ratio between the averages for the two parts and has been suggested as a measure of behavioral rigidity or inability to shift one's set.

The other "performance" test of perception was individually administered.

88-93. Thurstone's Color Form Movie. In 1952 Thurstone reported progress on an objective personality measure based on perceptual phenomena. For many years various psychologists have suggested that individuals differ fundamentally in that some respond more to the color and others more to the form aspects of visual stimuli. This idea underlies the famous perceptual experiment of Hermann Rorschach with colored ink blots. The movie consists of film sequences in which a color response and a form response compete without the subject's awareness of the fact. The subject is asked to call out the direction toward which he saw the design (or moving "band") move. In addition to color and form there are scores for four direction tendencies in the subject's answers. Each person then earns a total of six scores on the color form movie.

There were thus altogether twenty-eight scores based on personality measures; twenty-one of these were based on the paper-and-pencil tests, the other seven were more objective measures. Of the total of twenty-eight scores, five reached an F ratio which exceeded the .01 level of statistical significance, and four more had F ratios beyond the .05 level.

Sensory and musical aptitude tests. Limitations of time prohibited the use of

TABLE 6. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS FOR SOME MEASURES OF PERSONALITY FOR N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

	Name of test	h^2	F	p	N dz	N mz	Time required
	Thurstone Temperament Schedule						
66.	Factor A, Active	67	3.01	.01	35	45	20
67.	Factor B, Vigorous	59	2,43	.01	35	45	20
68.	Factor C. Impulsive	46	1.84	.05	35	45	20
69.	Factor D, Dominant	20	1.25		35	45	20
70.	Factor E, Stable	31	1.45		35	45	20
71.	Factor F, Sociable	47	1.90	.05	25	45	20
72.	Factor G, Reflective	06	1.06		35	45	20
	Cattell Junior Personality Quiz						
73.	Factor A, Emotional sensitivity	03	1.03		36	45	50
74.	Factor·B, Nervous tension	52	2.08	.05	36	45	50
75.	Factor C, Neuroticism	69	3.20	01	36	45	50
76.	Factor D, Will control	47	1.87	.05	36	45	50
77.	Factor E, Impatient dominance	—	.93		36	45	50
78.	Factor F, Cyclothymia	23	1.30		36	45	50
79.	Factor G, Adventurous cyclothymia	—	.93		36	45	50
80.	Factor H, Socialized morale	06	1.06		36	45	50
81.	Factor I, Independent dominance	—	.97		36	45	50
82.	Factor J, Energetic conformity	35	1.54		36	45	50
83.	Factor K, Surgency	31	1.45		36	45	50
84.	Factor L, Intelligence	09	1.10		36	45	50
85.	*Maudsley Word Connection Test*	22	1.28		35	42	20
	Stroop Color Naming Test						
86.	Verbal form (words/dots)	34	1.52		33	45	10
87.	Written form (words/dots)	22	1.29		22	41	10
	Thurstone Color-Form Movie						
88.	Color response	40	1.61		35	42	35
89.	Form response	29	1.42		35	42	35
90.	Upward response	40	1.67		35	42	35
91.	Downward response	19	1.24		35	42	35
92.	Left response	65	2.88	.01	35	42	35
93.	Right response	58	2.40	.01	35	42	35

the complete batteries of the Seashore and Wing Musical Aptitude Tests. The IPAT Test of Musical Preference was constructed by Cattell to attempt to measure personality characteristics related to predominant mood and preferred emotional tone in music preferences. It was understood at the time of selection that music preferences are largely the result of environmental influences, but the reported relationship to mental illness (Cattell and Anderson, 1953) made it seem worth while to include this highly imaginative technique. The results indicate that this may be an area which is promising for further study. Not unexpected is the finding that acuity of hearing is highly heritable, but intriguing is the finding of a differential between the right and left ear. Of the total of twenty-four scores, four reached an F ratio beyond the .01 level, and five more beyond the .05 level.

DISCUSSION

It may be helpful to summarize the findings in a somewhat condensed form before a brief discussion of the results. Table 8 presents for six areas of psychological functioning the number of measures which gave an F ratio significant at the 1 per cent and at the 5 per cent level of significance. The six areas have been ordered according to the degree to which hereditary com-

TABLE 7. HERITABILITY INDICES, F RATIOS, AND SIGNIFICANCE LEVELS OF SOME MEASURES OF MUSICAL ABILITY AND AUDITORY ACUITY FOR N_{mz} IDENTICAL AND N_{dz} FRATERNAL TWIN PAIRS

	Name of test	h^2	F	p	N dz	N mz	Time required
	Seashore						Mins.
94.	Pitch	—	.95		33	43	15
95.	Loudness	44	1.78		33	45	15
96.	Rhythm	52	2.07	.05	33	45	15
	Wing						
97.	Pitch change	12	1.14		34	44	15
98.	Memory	42	1.72	.05	34	43	15
	IPAT Test of Musical Preference (Cattell)						
99.	Factor A	41	1.70	.05	32	44	
100.	Factor B	—	.92		32	44	
101.	Factor C	31	1.44		32	44	
102.	Factor D	49	1.94	.05	32	44	
103.	Factor E	—	.42		32	44	
104.	Factor F	24	1.31		32	44	30a
105.	Factor G	27	1.38		32	44	
106.	Factor H	00	1.00		32	44	
107.	Factor I	34	1.51		32	44	
108.	Factor J	38	1.62		32	44	
109.	Factor K	11	1.12		32	44	
	Audiogram						
110.	2048, left	50	2.01	.05	36	44	
111.	2048, right	73	3.73	.01	37	45	40
112.	2896, left	12	1.14		36	44	
113.	2896, right	67	3.12	.01	37	45	
114.	5792, left	44	1.80		36	44	
115.	5792, right	72	3.70	.01	37	45	
116.	11584, left	19	1.43		36	44	40
117.	11584, right	59	2.46	.01	37	45	

ponents appear to be of importance. The F ratio was the ratio between the within-pair variance for the fraternal twins to the within-pair variance for the identical twins. The six areas examined in this paper are:

1. Thurstone's Primary Mental Abilities: seventeen variables. These tests were developed so as to have low intercorrelations between the primary abilities.
2. Motor Skills Tests: fourteen variables.
3. Other cognitive and achievements tests: eighteen variables.
4. Perceptual Tests: seventeen variables.
5. Sensory and musical aptitude tests: twenty-four variables.
6. Personality Tests. twenty-eight variables.

It should be emphasized again that this categorization is somewhat arbitrary and that some tests might perhaps as well belong in another rubric.

Table 8 tends to support the idea that individual differences in the various areas of psychological functioning are not to be attributed in an equal degree to hereditary components. That the personality measures and the measures of musical ability and preference are least determined by hereditary limitations is a not unexpected finding for somewhat different reasons.

To take up the measures of musical ability and preference first: while it is rather widely believed that there is an hereditary factor in musical talent (Anastasi, 1958; Farnsworth, 1958; Scheinfeld, 1956) it is equally widely realized that our present tests are not very adequate measures of this highly

TABLE 8. NUMBER OF MEASURES IN EACH OF SIX PSYCHOLOGICAL AREAS WHICH GAVE EVIDENCE OF A HEREDITARY COMPONENT AT ONE PERCENT AND FIVE PER CENT LEVELS OF SIGNIFICANCE

Area measures	No. of measures	p Value .01	.05	Not significant	
a. Primary mental abilities subtests	17	3	8	6	35.3
b. Motor skills	14	6	2	6	42.9
c. Perceptual tests	16	3	5	8	50.0
d. Cognitive and achievement tests	18	3	4	11	61.1
e. Sensory and musical	24	4	5	15	62,5
f. Personality test scores	28	5	4	19	67.9
All measures	117	24	28	64	54.7

complex phenomenon. It may be that only the exceptional talent of great composers and musicians has an hereditary factor and that such heights of skill have about as much to do with these tests as would the formulation of the theory of relativity with number ability as measured by tests of additions and multiplications. Nevertheless, a certain minimum level of proficiency on the elementary skills may have been a necessary, although not sufficient, condition for the personal development of the individual exceptional talent, since failure might have discouraged further striving in the area. It is also possible that limitations or defects are more properly the object of concern of hereditary studies, rather than proficiencies.

To turn to personality: it seems fairly sure that our present day measures of personality are inadequate for studies of hereditary factors. There is good reason to believe that efforts at clarification should be directed by studies of the heritability of scores on a wide variety of tests rather than by the success of such tests in differentiating various largely socially determined classes of persons, such as those engaged in different occupations, or different courses of study, or of persons affected by various mental illnesses or delinquent in one way or another. Yet such criteria are rightly considered of supreme importance in the development of better personality measures for diagnostic or predictive purposes. While work in these "applied" areas can be a source of fruitful ideas, the main effort in studies of hereditary factors in personality should be made with an open-minded broad search of all possible test instruments. That in this study cognitive and motor skills show somewhat more frequently evidence of hereditary influence fits in with the fact that (1) we can measure them better; (2) rudimentary differences can be noted in these domains even in babies; (3) some types of defects in these areas are known to be due to hereditary defects or antenatal and perinatal injuries.

In this connection it seems likely that studies of hereditary factors in personality would be more successful if such studies used observations on variables related to differences in infant personalities as their starting point. With proper attention to environmental conditions it may be that sex differences and family resemblance will provide better criteria for selection of variables to study than the social criteria mentioned here.

Before concluding, it is of interest to review briefly the measures which gave the clearest evidence of a hereditary contribution to the total variability. Table 9 lists the variables with an F ratio above 2.00 in order of the magnitude of

TABLE 9. VARIABLES WITH AN F RATIO BETWEEN DZ AND MZ WITHIN PAIR VARIANCE ABOVE 2.00 IN ORDER OF MAGNITUDE OF THE F RATIOS

* Audiogram 2048, right ear	3.73
* Audiogram 5792, right ear	3.70
* Card Sorting, left hand	3.42
* Tweezer Dexterity, right hand	3.40
* Mirror Drawing, right hand	3.38
Neuroticism, Cattell Junior Personality Quiz	3.20
* Audiogram 2896, right ear	3.12
Concealed Figures, Thurstone	3.05
Active, Thurstone Temperament Schedule	3.01
PMA Verbal, completion	3.01
Left response, Thurstone Color Form Movie	2.88
* Tweezer Dexterity, left hand	2.73
* Card Sorting, right hand	2.57
Kohs' Blocks	2.53
Stanford Spelling, from dictation	2.47
* Audiogram, right ear	2.46
Vigorous, Thurstone Temperament Schedule	2.43
Right Response, Thurstone Color Form Movie	2.40
* Santa Ana Dexterity, right hand	2.40
PMA Number, Addition	2.40
Gestalt Completion B, Thurstone	2.33
PMA Word Fluency, Suffixes	2.31
McGill Closure, Mooney	2.31
Heim AH4, Part 1	2.20

*A word of caution is in order with respect to the audiogram and motor skill tests. The possibility that an hereditary right-left preference or superiority is confounding the measurement of skill is a very real one. In future experiments it would be best to exclude individuals who display a noticeable advantage for the side not commonly superior in most subjects or to classify all individuals beforehand as lefthanded or righthanded and, independently, as preferentially using their left or right ear and then using the preferred side whether right or left.

the F ratios. In studying these variables with high F ratios, indicative of a significant hereditary contribution to the individual differences in this sample of high school age twins, it is first of all clear that a variety of distinct psychological attributes were found to be thus singled out. Of the cognitive variables it is well to keep in mind that in the groups studied by Thurstone in developing these measures these "factors" were not related.

It is the object of a companion paper to reconsider our data to see if there is evidence whether these measures are indeed unrelated in the subjects studied here and, more importantly, whether the hereditary components of variance are unrelated; by first asking the question whether there are significant correlations between the identical twin differences (due to environment) on various tests and then asking whether the correlations for the same test scores of the fraternal twin differences are greater or smaller than those for the score differences of the identical twins (thus indicating the presence or absence of association or identity of the hereditary components in these variables). We report here only that the correlations between scores (for all the subjects treated as single individuals) are generally low, and that the twin difference correlations present different patterns for the identical than for the fraternal twins for whom these correlations are sometimes significantly larger.

It is this author's conviction that just as with the personality variables, the usual consideration of the merits of cognitive tests which derive from significant social relevance for prediction of success or differentiation may have to be augmented by studies of perhaps less "successful" tests if we are to make progress in the study of hereditary components in psychological traits. This conviction arises from the facts (1) that the format of several tests of similar intent and of similar correlation with such criteria made an important difference in the F ratio and (2) that significant F ratios were found for tests which have only minor or nonexistent usefulness in present day applied psychological usage.

FINAL STATEMENT

The results reported indicate that hereditary factors play a role in many areas of human skilled performances, often in spite of the fact that these skills are highly practiced. Interesting discrepancies are indicative of challenging research possibilities which may lead to increased understanding of the complimentary roles of heredity and environment in the production of the individual differences observable in these tests. A more detailed and theoretical discussion of these results is reserved for a later paper, in which the findings will be compared with results obtained in studies by other investigators, and the question considered of the extent to which these variables may be under independent genetic control. In the meantime the curious reader is urged to consult the recent book on the genetics of behavior by Fuller and Thompson (1960).

SUMMARY

A battery of psychological tests furnishing one hundred seventeen separate scores was administered to forty-five pairs of identical twins and thirty-seven pairs of fraternal twins drawn from high schools in Ann Arbor, Ypsilanti, Dearborn, and Detroit, Michigan. Six groups of variables were measured: (1) Thurstone's Primary Mental Abilities Tests, (2) a variety of other cognitive and achievement tests, (3) motor skills, (4) perceptual skills, (5) sensory tests and measures of musical ability and interest, and (6) personality.

Each of these six groups contained a number of variables for which a significant degree of hereditary control was found, as indicated by an F above 2.00 for the ratio between the fraternal and identical within-pair variances. The proportion of variables for which at least a 5 per cent level of significance was reached varied between these six groups of variables. For 24, or 20.5 per cent of the one hundred seventeen variables, the F ratio exceeded the 1 per cent level of significance, and for another twenty-eight or 23.9 per cent the F ratio exceeded the 5 per cent level. Some evidence is presented which indicates that psychological tests which work well as predictors of social criteria such as success in academic studies, etc., may not be most promising for studies of hereditary factors.

ACKNOWLEDGMENTS

Invaluable advice on the selection of tests was received from E. Lowell Kelly, Raymond B. Cattell, the late L. L. Thurstone, and James R. Birren. Benjamin White participated in the initial planning of the study. The skillfull assistance of Mrs. Lin Wong, Mrs.

Louella Goodall, Mrs. Phyllis Kao, Mrs. Felicia Lewis, and above all, of Mrs. Lillian Kelly is gratefully remembered. Mrs. Elsie Long patiently typed and retyped the manuscript. This investigation would have been impossible without the cooperation of the schools in Ann Arbor, Ypsilanti, Dearborn, and Detroit. Thanks are due to all the administrators, teachers, and twins for their wholehearted cooperation.

REFERENCES

ALEXANDER, W. P. 1935. Intelligence, concrete and abstract: a study of differential traits. *Brit. J. Psychol. Monogr. Suppl.* Vol. 6, No. 19. London: Cambridge Univ. Press.

AMMONS, R. B., AND HUTH, R. W. 1949-a. The Full-Range Picture Vocabulary Test I. Preliminary Scale. *J. Psychol.* 28:51-64.

AMMONS, R. B., AND HOLMS, J. C. 1949-b. The Full-Range Picture Vocabulary Test: III. Results for a preschool-age population. *Child Develpm.* 20:5-14.

AMMONS, R. B., AND RACHIELE, L.D. 1950-a. The Full-Range Picture Vocabulary Test II. Selection of items for final scales. *Educ. Psychol. Measmt.* 10: 307-319.

AMMONS, R. B., ARNOLD, P. R., AND HERRMANN, R. S. 1950-b. The Full-Range Picture Vocabulary Test: IV. Results for a white school population. *J. Clin. Psychol.* 6: 164-169.

AMMONS, R. B., LARSON, W. L., AND SHEARN, C. R. 1950-c. The Full-Range Picture Vocabulary Test: V. Results of an adult population. *J. Consult. Psychol.* 14: 150-155.

ANASTASIA, A. 1958. *Differential psychology.* New York: Macmillan.

ARTHUR, G. 1944. A non-verbal test of logical thinking. *J. Consult. Psychol.* 8: 33-34.

BENNETT, G. K., SEASHORE, H. G., AND WESMAN, A. G. 1947. *Differential aptitude test.* New York: Psychological Corp.

CATTELL, R. B., AND ANDERSON, J. C. 1953. The measurement of personality and behavior disorders by the I.P.A.T. Music Test. *J. Appl. Psychol.* 37: 446-454.

CATTELL, R. B., BELOFF, J., AND BELOFF, H. 1950. *The junior personality quiz.* Champaign, Ill.; Institute for Personality and Ability Testing.

CLARK, P. J. 1956. The heritability of certain anthropometric characters as ascertained from measurements of twins. *Amer. J. Hum. Genet.* 8: 49-54.

CORNELL, E., AND COXE, W. W. 1934. *A performance ability scale: examination manual,* Yonkers-on-Hudson, N. Y.: World Book Co.

CROWN, S. 1952. The word connection list as a diagnostic test: norms and validations. *Brit. J. Psychol.* 43: 103-112.

DUBOIS, P. H., AND GLESER, Goldine 1948. The object-aperture test—a measure involving visualization in three dimensions (abstr.). *Amer. Psychologist* 3: 363.

EYSENCK, H. J. 1952. *The scientific study of personality.* London: Routledge and Kegan Paul.

FARNSWORTH, P. R. 1958. *The social psychology of music.* New York: Dryden Press.

FRENCH, J. W. 1951. *The description of aptitude and achievement tests in terms of rotated factors.* Psychometric Monograph 5. Chicago: Univ. Chicago Press.

FULLER, J., AND THOMPSON, W. R. 1960. *Behavior genetics.* New York: Wiley.

GRAY, W. S. 1915. *Standardized oral reading paragraphs test.* Bloomington, Ill.: Public School Publishing Co.

GRAY, W. S. 1920. *Studies of elementary school reading through standardized tests.* Suppl. Educ., Monogr. No. 1. Chicago: Univ. Chicago Press.

GUILFORD, J. P. 1956. The structure of intellect. *Psychol. Bull.* 53: 287-293.

GUILFORD, J. P. 1960. Morphological model of human intelligence. *Science* 131: 1318.

HEIM, A. 1947. An attempt to test high grade intelligence. *Brit. J. Psychol.* 37: 70-81.

HEIM, A. 1955. *Test AH4: group test of intelligence.* London: National Foundation for Educational Research.

HUNT, T. 1928. The measurement of social intelligence. *J. Appl. Psychol.* 12: 317-334.

KELLEY, T. L., RUCH, G. M., AND TERMAN, L. M. 1940. *Stanford Achievement Tests, Form D.* Yonkers-on-Hudson, N. Y.: World Book Co.

KOHS, S. C. 1923. *Intelligence measurement, a psychological and statistical study based upon the Block-design Tests.* New York: Macmillan.

LOEVINGER, J. 1951. Intelligence. In *Theoretical foundation of psychology*, H. Nelson, ed. New York: Van Nostrand, pp. 557-603.

MACKINNON, D. W., AND HENLE, M. 1948. *Experimental studies in psycho-dynamics.* Cambridge: Harvard Univ. Press.

MELTON, A. W. (ed.) 1947. *Apparatus tests.* Washington: Army Air Force Psychol. Program Res. Rep. No. 4.

MILES, W. R. 1950. Static equilibrium. In *Methods in medical research,* R. W. Gerard, ed. 3: 157-165. Chicago: Yearbook Publ.

MILLER, D. R. 1945. *Survey of mechanical insight.* Los Angeles: California Test Bureau.

MOONEY, C. M. 1954. A factorial study of closure. *Canad. J. Psychol.* 8:51-60.

MOSS, F. A., HUNT, T., AND OMWAKE, K. T. 1930. *Social intelligence test.* Washington: George Washington Univ.

O'CONNOR, J. 1928. *Born that way.* Baltimore: Williams and Wilkins.

O'CONNOR, J. 1938. *Administration and norms for the finger dexterity test and the tweezer dexterity test.* Boston: Technical Report No. 16, Human Engineering Laboratory.

RAVEN, J. C. 1938. *Progressive matrices, test and instructions.* London: H. K. Lewis.

RAVEN, J. C. 1941. Standardization of progressive matrices. *Brit. J. Med. Psychol.* 19: 137-150.

SCHEINFELD, A. 1956. *The human heredity handbook.* New York: J. P. Lippincott Co.

SUTTON, H. E., VANDENBERG, S. G., AND CLARK, P. J. 1962. The hereditary abilities study: selection of twins, diagnosis of zygosity and program of measurements. *Am. J. Hum. Genet.* 14: 52-63.

THURSTONE, L. L. 1938. *Primary mental abilities.* Psychometric Monograph No. 1. Chicago: Univ. Chicago Press.

THURSTONE, L. L., AND THURSTONE, T. 1941-a. *The Chicago tests of primary mental abilities.* Chicago, Ill, American Council on Education (distributed by Science Research Associates).

THURSTONE, L. L., AND THURSTONE, T. 1941-b. *Factorial studies of intelligence.* Psychometric Monograph No. 2. Chicago: Univ. Chicago Press.

THURSTONE, L. L. 1949. *A factorial study of perception.* Psychometric Monograph No. 4. Chicago: Univ. of Chicago Press.

THURSTONE, L. L. 1952. *Progress report on a color-form test.* Chicago: The Psychometric Laboratory.

WECHSLER, D. 1949. *Wechsler intelligence scale for children.* New York: Psychological Corp.

WECHSLER, D. 1944. *The measurement of adult intelligence.* New York: Psychological Corp.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISNER, P. B., AND WAPNER, S. 1954. *Personality through perception.* New York: Harper.